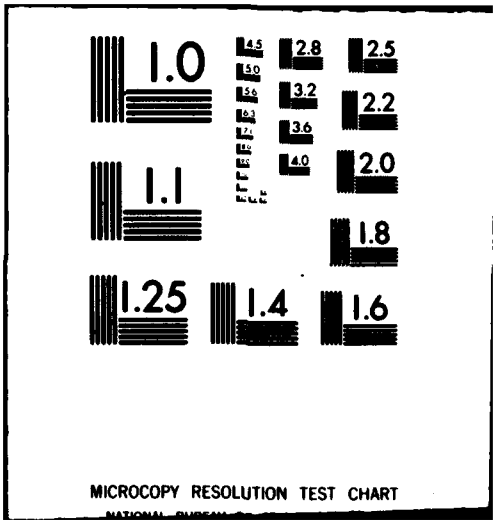
AD-A091 035 DELAWARE UNIV NEWARK APPLIED MATHEMATICS INST F/6 12/1
AN ILLUSTRATIVE PROBLEM IN COMPUTATIONAL PROBABILITY.(U)
JUN 80 M F NEUTS AFOSR-77-3236
UNCLASSIFIED TR-608 AFOSR-TR-80-0987 NL

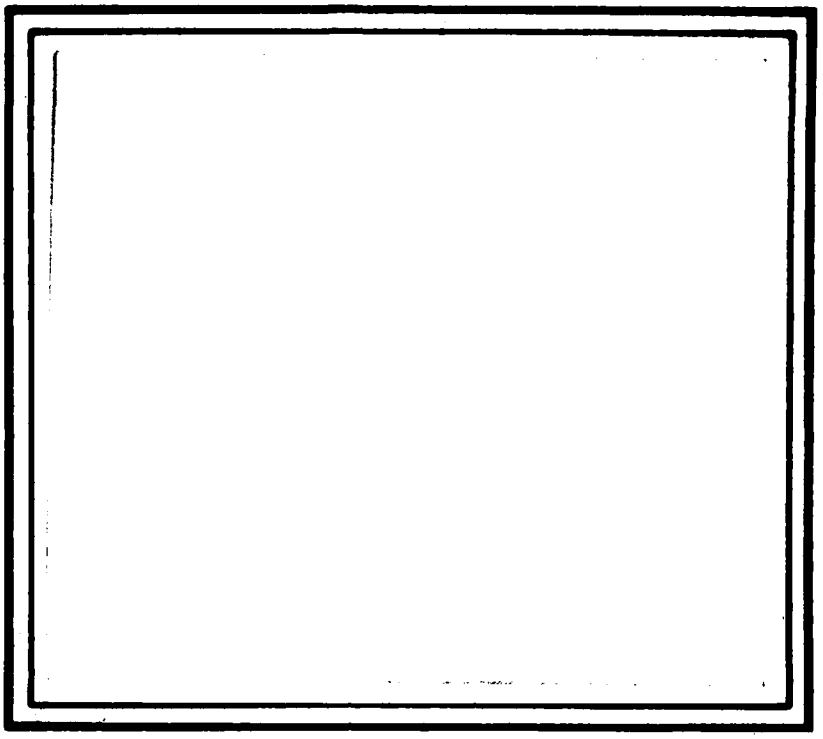
END
DATE FILMED
12-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 80-0987

(5) LEVEL II

AD A091035



See 14/73

DTIC
ELECTE
OCT 3 1 1980
S D
B

DDC FILE COPY.

APPLIED MATHEMATICS INSTITUTE

University of Delaware
Newark, Delaware

80 10 6 09

Approved for public release;
distribution unlimited.

(5) LEVEL II

An Illustrative Problem in Computational Probability

by

Marcel F. Neuts
Department of Mathematical Sciences
University of Delaware
Newark, Delaware 19711

Applied Mathematics Institute
Technical Report # 60B
June 1980

This paper was prepared as a contribution to the Souvenir Volume of the First Conference of the Indian Society for the Theory of Probability and its Applications, held at Tirupati, July 8 - 10, 1980.

This research was supported by the National Science Foundation under Grant No. ENG-7908351 and by the Air Force Office of Scientific Research under Grant No. AFOSR 77-3236.

ABSTRACT

An elementary, but useful problem in reliability design is used to illustrate some of the principles and concerns of computational probability. A number n of items with exponential life times are placed in parallel. The time until all items have failed is denoted by U_n. It is desirable that U_n exceed the duration T, where T is a random variable, independent of the n life times. We wish to determine the smallest value of n for which $P\{U_n > T\}$ exceeds $1 - \varepsilon$. It is shown that if T has a delayed distribution of phase type, this may be done by a recursive algorithm, which avoids numerical integrations.

KEY WORDS

Reliability, distributions of phase type, computational probability.

1. Formulation of the Problem

The concern for algorithmically tractable solutions to problems in probability adds an interesting new dimension to their analysis. In the construction of efficient and stable algorithms, certain principles, which have not been given much attention in the classical analytic presentations, need to be followed. In this paper, some of these principles are illustrated by a simple example from reliability theory.

We first consider n items with independent, identically distributed, exponential life times with parameter μ. These items are placed in parallel and system failure occurs at the time U_n when all n items have expired. The parallel system is to be used during a mission of duration T, where T is a random variable, independent of the life times of the items. The distribution of T is denoted by $F(\cdot)$.

The probability $q(n) = P\{U_n \leq T\}$, that the system fails before the end of the mission is readily given by

$$q(n) = \int_0^\infty (1-e^{-\mu x})^n \, dF(x) , \qquad \text{for } n \geq 1 . \tag{1}$$

Clearly $q(n)$ is strictly decreasing in n and tends to zero as n tends to infinity. For every $\varepsilon > 0$, the quantity $n^* = \min \{n: q(n) < \varepsilon\}$ is well-defined. n^* is the smallest number of items, which will yield a reliability of at least $1 - \varepsilon$.

The numerical computation of n^* requires some care. It would e.g. be inadvisable to rewrite (1) as

$$q(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(k\mu) , \tag{2}$$

where $f(\cdot)$ is the Laplace-Stieltjes transform of $F(\cdot)$. For large n, the

quantity $q(n)$ is given by (2) as the sum of terms of alternating sign. Numerical implementation of (2) usually suffers from loss of significance.

This serves to illustrate our first point. Mathematically equivalent solutions may be vastly different in their suitability for numerical computation. A number of the published analytic solutions to stochastic models are worthless for or hazardous in numerical implementation. Algorithmic tractability deserves to be included in technical discussions of probability models to a far greater extent than is presently done.

In evaluating the quantities $q(n)$ by numerical integration in Formula (1), considerable care is needed. For large values of n, the integrand is small for small to moderate x. The improper Riemann-Stieltjes integrals need to be truncated at a high value of x, so as to avoid neglecting a substantial tail contribution. For specific probability distributions $F(\cdot)$, we may often perform an appropriate change-of-variable, so that e.g. the interval of integration will become finite or so that a quadrature method may be applied. This, of course, depends on particular features of $F(\cdot)$.

In any case - and this is our second point - it will not be possible to write one computer code to handle a very wide variety of probability distributions $F(\cdot)$. Except for a small number of classical particular probability distributions, such as e.g. the gamma family, it is necessary to analyze the numerical integration procedure from first principles. This is not a matter of mere computer programming. For this - and *a fortiori* for more complex problems - this analysis should be the concern and responsibility of the probabilist.

In many problems, we may exploit particular probabilistic features to expedite the computations. For the versatile class of distributions of phase type, we shall show that the quantities $q(n)$ may be evaluated without numerical

integrations. Before doing so, we shall consider a more general version of the problem. Let there be n items originally, which fail according to a pure death process, i.e. the times between successive failures are independent and exponentially distributed with parameters $\mu_n, \ldots, \mu_1$. The time-to-failure then has a generalized Erlang distribution. If $\phi_k^{(n)}(t)$ denotes the probability that k of the n original items are still alive at time t, the quantity $q(n)$ is given by

$$q(n) = \int_0^\infty \phi_0^{(n)}(x)\, dF(x) , \qquad \text{for } n \geq 0 . \tag{3}$$

The earlier problem corresponds to the choice $\mu_j = j\mu$, for $1 \leq j \leq n$.

2. The Case Where $F(\cdot)$ is of Phase Type

$F(\cdot)$ is a distribution of phase type (PH) if and only if there is a finite-state Markov process, with one absorbing state and all other states transient, in which $F(\cdot)$ is the distribution of the time till absorption [1,2]. The generator Q of such a Markov process is then of the form

$$Q = \begin{vmatrix} T & \underline{T}^\circ \\ \underline{0} & 0 \end{vmatrix} ,$$

where T is a square matrix, say of order m, with negative diagonal elements and nonnegative off-diagonal elements. The matrix T is nonsingular and $T\,\underline{e} + \underline{T}^\circ = \underline{0}$. The initial probability vector is denoted by $[\underline{\alpha}, \alpha_{m+1}]$. The initial PH-distribution $F(\cdot)$ is then given by

$$F(x) = 1 - \underline{\alpha} \exp(Tx)\, \underline{e} , \qquad \text{for } x \geq 0 . \tag{4}$$

In the sequel, we shall assume for convenience that $\alpha_{m+1} = 0$, so that $F(\cdot)$ is continuous on $[0,\infty)$.

Let now $\psi_i(n)$ be the conditional probability that $U_n \leq T$, given that the Markov process Q starts in the state i. The m-vector with components $\psi_i(n)$, $1 \leq i \leq m$, is denoted by $\underline{\psi}(n)$. It is then clear that

$$q(n) = \underline{\alpha}\, \underline{\psi}(n) , \qquad \text{for } n \geq 0 . \tag{5}$$

The vectors $\underline{\psi}(j)$, $0 \leq j \leq n$, are recursively given by

$$\underline{\psi}(0) = \underline{e} , \tag{6}$$

$$\underline{\psi}(j) = \mu_j(\mu_j I - T)^{-1}\, \underline{\psi}(j-1) , \qquad \text{for } 1 \leq j \leq n .$$

It is clear that $\underline{\psi}(0) = \underline{e}$. Furthermore, by conditioning on the time of the first of the j failures, given that the system starts with $j \geq 1$ items, we obtain by using the Markov property that

$$\underline{\psi}(j) = \int_0^\infty \exp\,(Tx)\, e^{-\mu_j x}\, \mu_j\, dx \cdot \underline{\psi}(j-1) ,$$

which readily yields (6).

From Formulas (5) and (6), we see that the $q(n)$ may be computed by the numerical solution of systems of linear equations. We evaluate the successive vectors $\underline{\psi}(j)$, $j \geq 1$, and stop as soon as $\underline{\alpha}\, \underline{\psi}(j)$ becomes less than the prescribed ε. The integer n^* is then given by the index j at which computation halts.

The numerical solution of the linear equations

$$\underline{\psi}(j) - \mu_j^{-1}\, T\, \underline{\psi}(j) = \underline{\psi}(j-1) , \tag{7}$$

is highly stable. In particular, they are ideally suited for Gauss-Seidel iteration. As we shall discuss below, there may be some merit in using this procedure, although it is somewhat slower than direct Gauss elimination.

3. The Case Where $F(\cdot)$ is a Delayed PH-distribution.

In practice, it is unrealistic to assume that the distribution $F(\cdot)$ has support on $(0,\infty)$. The duration T will usually exceed some fixed value $a > 0$ with probability one. It is easy to modify the preceding algorithm to handle this case also. Let $F(\cdot)$ be of the form

$$F(x) = 0 , \qquad \text{for } x \leq a ,$$

$$= 1 - \underline{\alpha} \exp [T(x-a)] \underline{e} , \qquad \text{for } x \geq a .$$

It may be called a delayed PH-distribution.

The quantity $q(n)$ is then given

$$q(n) = \sum_{k=0}^{n} \phi_k^{(n)}(a)\, \hat{q}(k) , \qquad \text{for } n \geq 0 , \tag{8}$$

where the $\hat{q}(k)$ are the quantities, given by Formula (5) . In the particular model, with which we started this discussion, Formula (8) becomes

$$q(n) = \sum_{k=0}^{n} \binom{n}{k} \left(1-e^{-\mu a}\right)^{n-k} e^{-k\mu a}\, \hat{q}(k) , \qquad \text{for } n \geq 0 .$$

The binomial probabilities $\phi_k^{(n)}(a)$ are easily evaluated. In general, the probabilities $\phi_k^{(n)}(t)$ may be computed by the numerical solution of the simple differential equations

$$\begin{aligned} \frac{d}{dt}\phi_n^{(n)}(t) &= -\mu_n \phi_n^{(n)}(t) , \\ \frac{d}{dt}\phi_k^{(n)}(t) &= -\mu_k \phi_k^{(n)}(t) + \mu_{k+1} \phi_{k+1}^{(n)}(t) , \qquad \text{for } 1 \leq k < n , \\ \frac{d}{dt}\phi_0^{(n)}(t) &= \mu_1 \phi_1^{(n)}(t) , \end{aligned} \tag{9}$$

for $t \geq 0$, with the initial conditions $\phi_n^{(n)}(0) = 1$, $\phi_k^{(n)}(0) = 0$, for $0 \leq k < n$.

We believe that the preceding discussion illustrates the point that for many problems of practical interest, probabilistic insight may lead to natural and efficacious algorithms. There is one additional point to be made in conclusion. Our simple example again serves to illustrate that point well.

In practical situations, we often wish to vary parameters of the model in a systematic manner. It is desirable to "nest" the successive implementations of the basic algorithm, so that quantities computed in earlier runs may be reused or efficiently updated.

If, for example, we wish to vary the parameter a in Formula (8), it suffices to start with the smallest value of a and to integrate the differential equations (9) progressively up to the largest desired value of a. The same sequence $q(k)$, $k \geq 0$, is used throughout.

If we wish to modify one or more elements of the matrix T, we can solve the equations (7) by an iterative method and use the preceding solution vectors as starting solutions.

These considerations are particularly important when the actual computations are carried out in the conversational mode at a remote terminal. The algorithm then becomes a powerful tool, which may be "interrogated" to assess the effect of varying the parameters of the model. In complex models, it is mostly impossible to do this by analytic methods only. The times between the successive printouts of results are usually a dead loss to the user and it is desirable to keep these as short as possible.

Once the algorithmic analysis of one stochastic model has been carried out, it often provides insights into the analysis of related models. The preceding discussion may easily be modified to handle the case where system

failure occurs when only r out of the original n items survive. The times between successive failures may also be chosen to have distributions of phase type instead of the (particular) exponential distributions, used in our discussion. The algorithm then becomes more complex, but the underlying mathematical ideas remain the same. To identify the structural properties, useful in the construction of efficient algorithms, is the proper objective of computational probability.

References

[1] Neuts, M. F. (1975)
Probability Distributions of Phase Type
in Liber Amicorum Professor Emeritus H. Florin,
Univ. of Louvain, Belgium, 173-206.

[2] Neuts, M. F. (1981)
Matrix-geometric Solutions in Stochastic Models - An Algorithmic Approach
The Johns Hopkins University Press, Baltimore, MD, to appear.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

AFOSR-TR-80-0987	2. GOVT ACCESSION NO. AD-A091035
4. TITLE (and Subtitle) AN ILLUSTRATIVE PROBLEM IN COMPUTATIONAL PROBABILITY.	5. TYPE OF REPORT & PERIOD COVERED Interim rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Marcel F. Neuts	8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3236, NSF-ENG79-08351
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Delaware Applied Mathematics Institute Newark, DE 19711	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	12. REPORT DATE Jun 1980 13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reliability, distributions of phase type, computational probability

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An elementary, but useful problem in relability design is used to illustrate some of the principles and concerns of computational probability. A number n of items with exponential life times are placed in parallel. The time until all items have failed is denoted by U_n. It is desirable that U_n exceed the duraction T, where T is a random variable, independent of the n^{th} life times. We wish to determine the smalles value of n for which $P(U_n > T)$ exceeds $1 - \varepsilon$. It is shown that if T has a delayed distribution of phase type, this may be done by a recursive algorithm, which avoids numerical integrations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

411278